

---

# Peptide Models XVI. The Identification of Selected HCO–L–SER–NH<sub>2</sub> Conformers via a Systematic Grid Search Using *Ab Initio* Potential Energy Surfaces

---

ANDRÁS PERCZEL,<sup>a,b</sup> ÖDÖN FARKAS,<sup>b,c</sup> and IMRE G. CSIZMADIA<sup>b,\*</sup>

<sup>a</sup> Department of Biochemistry, University of Oxford, Oxford OX1 2QU, United Kingdom

<sup>b</sup> Institute of Organic Chemistry, Eötvös University Budapest, Budapest, Hungary

<sup>c</sup> Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 1A1 Canada

Received 4 January 1994; accepted 13 June 1995

---

## ABSTRACT

Multidimensional conformational analysis (MDCA) predicted the existence of nine stable backbone conformations ( $\alpha_L$ ,  $\alpha_D$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\gamma_D$ ,  $\delta_L$ ,  $\delta_D$ ,  $\epsilon_L$ , and  $\epsilon_D$ ) on the 2D-Ramachandran map,  $E = E(\phi, \psi)$ , for a single amino acid diamide (HCONH–CHR–CONH<sub>2</sub>). The potential energy hypersurfaces ( $E = E[\phi, \psi, \chi_1, \chi_2]$ ) of For–L–Ser–NH<sub>2</sub> associated with the  $\alpha_L$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\delta_L$ , and  $\epsilon_L$ -type stable backbone orientations are investigated in this article. An appropriate number of side-chain rotamers is associated with each of the backbone conformers. In the case of serine, where  $R = -CH_2OH$ , the two sidechain torsional angles ( $\chi_1, \chi_2$ ) should lead to  $3 \times 3 = 9$  different sidechain orientations according to MDCA. For certain backbone structures, some of the sidechain conformations were nonexistent. © 1996 by John Wiley & Sons, Inc.

---

## Introduction

Recent advances in genetic engineering provide an indispensable tool to replace certain amino acids in proteins in order to find new enzyme inhibitors, more selective drugs, or biologi-

\* Author to whom correspondence should be addressed.

cally important ligand molecules. These point mutation experiments involve the exchange of a given amino acid by a more suitable candidate located at a targeted sequential position. Such a replacement of the sidechain functional group affects the conformation of the neighboring amino acid residues, often influencing the conformation of the macromolecule. Basically, the backbone–backbone and the backbone–sidechain interactions affect the rel-

ative orientation of the constituent amino acid residues, thus altering the global structure of the macromolecule. However, little is known about the conformational integrity involved of the associated folding process.

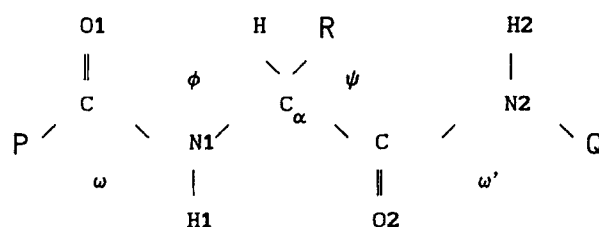
Even with high-field nuclear magnetic resonance (NMR) spectrometers, the structure analyses of flexible amide systems are almost impossible because multiple conformers are undergoing structural conversions with a time scale typically faster than those resolved by NMR.<sup>1-5</sup> Furthermore, the NMR-determined interproton distances (NOE, or nuclear Overhauser effect, results<sup>6</sup>) are associated with an averaged structure usually composed of several different conformers. Moreover, the explicit advanced knowledge of the conformational data (bond lengths, bond angles, and torsional angles) is indispensable for quantitative structure determinations performed in solution by NMR spectroscopy. The assignment and interpretation of the normal mode vibrations recorded by an infrared (IR) spectrometer also require the knowledge for the individual conformers. Thus advanced knowledge of the geometry of the conformers that are required for spectroscopical data analyses can now be investigated at least for small peptides by *ab initio* computational methods. *Ab initio* calculations yield results which are much more adequate than those obtained previously using parameterized molecular mechanics (MM),<sup>7-9</sup> molecular dynamics,<sup>10,11</sup> and semiempirical molecular orbital, (MO) computations.<sup>12</sup> Structures of the individual conformers can be determined with a higher degree of confidence by *ab initio* type calculations,<sup>13,29</sup> which are basically free from empirical parameters.

The conformation analyses of amino acid diamides (R'CONH-CHR-CNOR''') as the building units of proteins are of continuous interest. For single amino acid diamides, a maximum of nine different backbone conformations is predicted on the basis of theoretical conformation analyses.<sup>21,30,31</sup> The systematic conformation analyses of For-Gly-NH<sub>2</sub> (ref. 19-22), For-L-Ala-NH<sub>2</sub> (ref. 19-22), For-D-Ala-NH<sub>2</sub> (ref. 21), and For-L-Val-NH<sub>2</sub> (ref. 22-23) derivatives were previously studied. Serine diamide (frequently found in proteins) is the next logical candidate for conformational analysis, which is the subject of this study. The importance of serine may be underlined by the fact that the hydroxymethyl group (the sidechain of serine) can be involved directly in enzymatic reactions, as demonstrated for proteins of the serine protease family (trypsin,

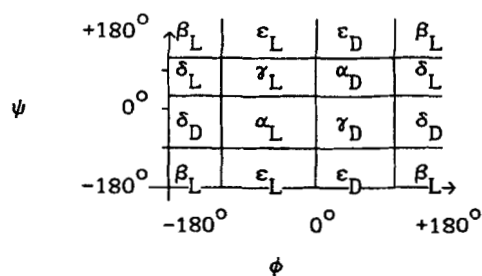
chymotrypsine, etc.), or can have an indirect influence on proteins by stabilizing different secondary structural elements. Serine is often found in the  $\alpha$ -helical region of globular proteins as well as in the  $(i + 2)$  position of a  $\beta$ -turn, where a backbone/sidechain stabilizing effect is expected to be operative. The lack of the  $\alpha_L$ -type minima, which are the conformational building blocks of the right-handed  $\alpha$ -helix, is consistently observed for *ab initio* type structure investigations of For-L-Xxx-NH<sub>2</sub> systems, where Xxx = Gly, Ala, and Val). Moreover, the restabilization of the annihilated  $\alpha_L$  backbone conformation has already been demonstrated.<sup>32</sup>

Multidimensional conformational analysis<sup>30,31</sup> predicts the existence of nine minima<sup>21</sup> on the full Ramachandran<sup>33</sup> map ( $E = E[\omega, \phi, \psi, \omega']$ ), where the torsional angles are defined according to IUPAC-IUB convention, as shown in Scheme 1. Since  $\omega$  and  $\omega'$  have torsional angle values typically 180° for trans peptide bonds (and 0° for cis peptides), the energy expression can be simplified as  $E_{\omega=\omega'=\text{const.}} = E[\phi, \psi]$ , where energy is dependent on only two torsional variables. Protein chemists have accepted the IUPAC-IUB convention by varying both the  $\phi$  and the  $\psi$  in between -180° and +180°. From a conformational perspective, it is more convenient to identify minima within their catchment region by using the  $0^\circ \leq \phi \leq 360^\circ$  and  $0^\circ \leq \psi \leq 360^\circ$  cut of the  $[\phi, \psi]$  space because such a cut does not fragment any of the catchment regions. Scheme 2 illustrates the nine minima in an idealized fashion. Note that the top representation is drawn according to the IUPAC-IUB<sup>34a</sup> convention. The bottom representation illustrates the topologically more useful cut as introduced earlier.<sup>21-22</sup>

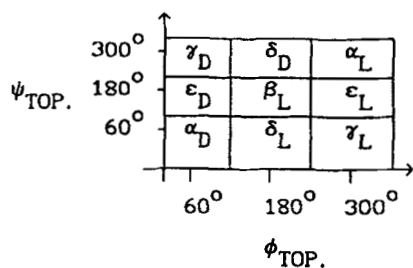
A preliminary database search was performed using the Brookhaven X-ray structure data<sup>34b</sup> for 78 selected proteins with relatively good R factor and resolution lower than 2.5 Å. The analyses of serine backbone conformation demonstrated the dominance (93.1%) of the  $\alpha_L$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\delta_L$ , and  $\epsilon_L$



**SCHEME 1.** Torsional angles ( $\omega$ ,  $\phi$ ,  $\psi$ ,  $\omega'$ ) of an N- and C-protected amino acid residue.



IUPAC-IUB representation



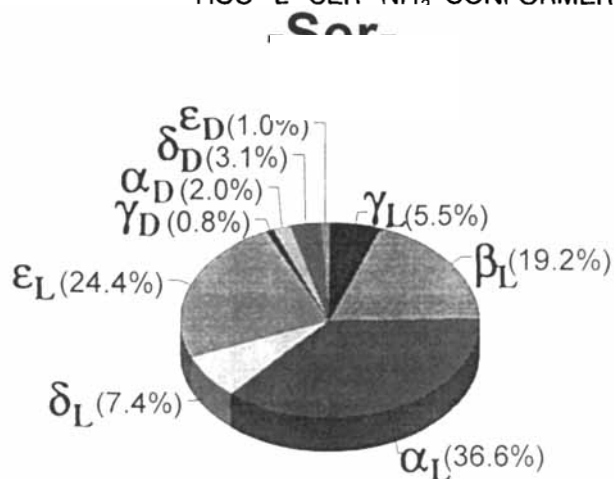
Topology oriented representation

**SCHEME 2.** Two representations of the Ramachandran basis ( $\phi$ ,  $\psi$ ).

(all five with L subscript) structures, while the remaining four backbone conformations  $\alpha_D$ ,  $\gamma_D$ ,  $\delta_D$ , and  $\varepsilon_D$  (all with D subscript) added up to a sum of 6.9% (Fig. 1). The five most important conformations were assigned on the basis of previous *ab initio* calculations<sup>28</sup> using the  $\alpha_L$  ( $\phi_{\alpha_L} = -68.6^\circ$ ,  $\psi_{\alpha_L} = -17.5^\circ$ ),  $\beta_L$  ( $\phi_{\beta_L} = -167.6^\circ$ ,  $\psi_{\beta_L} = +169.9^\circ$ ),  $\gamma_L$  ( $\phi_{\gamma_L} = -84.4^\circ$ ,  $\psi_{\gamma_L} = +68.7^\circ$ ),  $\delta_L$  ( $\phi_{\delta_L} = -126.2^\circ$ ,  $\psi_{\delta_L} = 26.5^\circ$ ), and  $\varepsilon_L$  ( $\phi_{\varepsilon_L} = -74.7^\circ$ ,  $\psi_{\varepsilon_L} = +167.8^\circ$ ) conformational centers. Therefore, the detailed grid search for all the possible For-L-Ser-NH<sub>2</sub> conformers was started with sidechain maps ( $f[\chi_1, \chi_2]$ ) associated with the  $\alpha_L$  (ref. 32),  $\beta_L$ ,  $\gamma_L$  (ref. 35),  $\delta_L$ , and  $\varepsilon_L$  backbone types.

## Computational Methods

The conformations of the For-L-Ser-NH<sub>2</sub> geometries were determined using  $(3n - 6)$  natural internal coordinates<sup>36,37</sup> in total geometry optimization. Approximate locations of backbone minima were determined by using semiempirical (AM1/MOPAC)<sup>12</sup> and/or molecular mechanic (PCMODEL, ECEPP/2)<sup>7,8</sup> methods. Fully relaxed *ab initio* calculations were performed to determine sequentially the optimized  $[\phi, \psi]$  positions associated with different sidechain orientations.

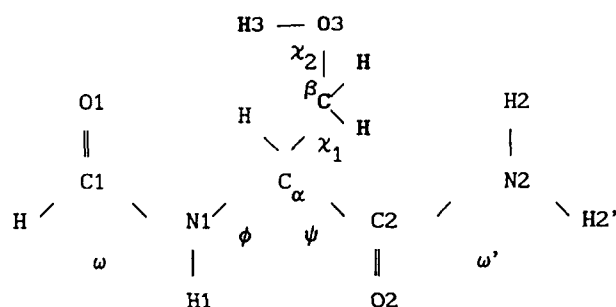
**FIGURE 1.** Analysis of serine backbone conformations

**FIGURE 1.** Analysis of serine backbone conformations in 78 selected proteins using the Brookhaven X-ray structure data.<sup>34b</sup> The dominance (subtotal 93.1%) of the  $\alpha_L$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\delta_L$ , and  $\varepsilon_L$  (all five with L subscript) structures is significant. The five conformations with capitalized L were assigned on the basis of previous *ab initio* calculations<sup>28</sup> using the  $\alpha_L$  ( $\phi_{\alpha_L} = -68.6^\circ$ ,  $\psi_{\alpha_L} = -17.5^\circ$ ),  $\beta_L$  ( $\phi_{\beta_L} = -167.6^\circ$ ,  $\psi_{\beta_L} = +169.9^\circ$ ),  $\gamma_L$  ( $\phi_{\gamma_L} = -84.4^\circ$ ,  $\psi_{\gamma_L} = +68.7^\circ$ ),  $\delta_L$  ( $\phi_{\delta_L} = -126.2^\circ$ ,  $\psi_{\delta_L} = 26.5^\circ$ ), and  $\varepsilon_L$  ( $\phi_{\varepsilon_L} = -74.7^\circ$ ,  $\psi_{\varepsilon_L} = +167.8^\circ$ ) conformational centers.

Alternative sidechain geometries with a common backbone conformation (e.g., extended-like [ $\beta_L$ ] backbone structures) resulted in characteristic  $[\phi, \psi]$  values. At such critical points of the Ramachandran surface, sidechain maps ( $f[\chi_1, \chi_2]$ ) were generated using  $30^\circ$  increments along both dimensions. Thus, sidechain maps ( $f[\chi_1, \chi_2]$ ) associated with the  $\alpha_L$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\delta_L$ , and  $\varepsilon_L$  backbone conformations were determined, where each  $f[\chi_1, \chi_2]$  map consisted of 169 grid points. Since these maps were associated with a particularly fixed backbone conformation, they are called fixed sidechain maps ( $f_{bb=\text{fixed}}[\chi_1, \chi_2]$ ), even though  $[(3n - 6) - 4]$  internal coordinates were fully relaxed. Minimization was performed using a standard 3-21G basis set on a modified GDIIIS<sup>38</sup> and the EFGDIIS<sup>39</sup> gradient method, and energy was calculated by using the TX90 program<sup>40</sup> within a modified MOPAC 5.00.

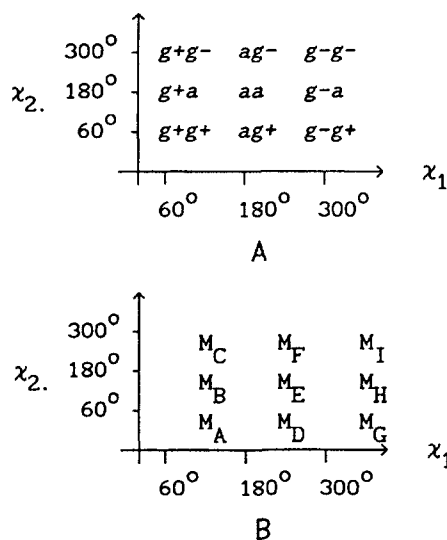
## Results and Discussion

The hydroxymethyl group (HO-CH<sub>2</sub>-) of the serine residue (Scheme 3) is expected to

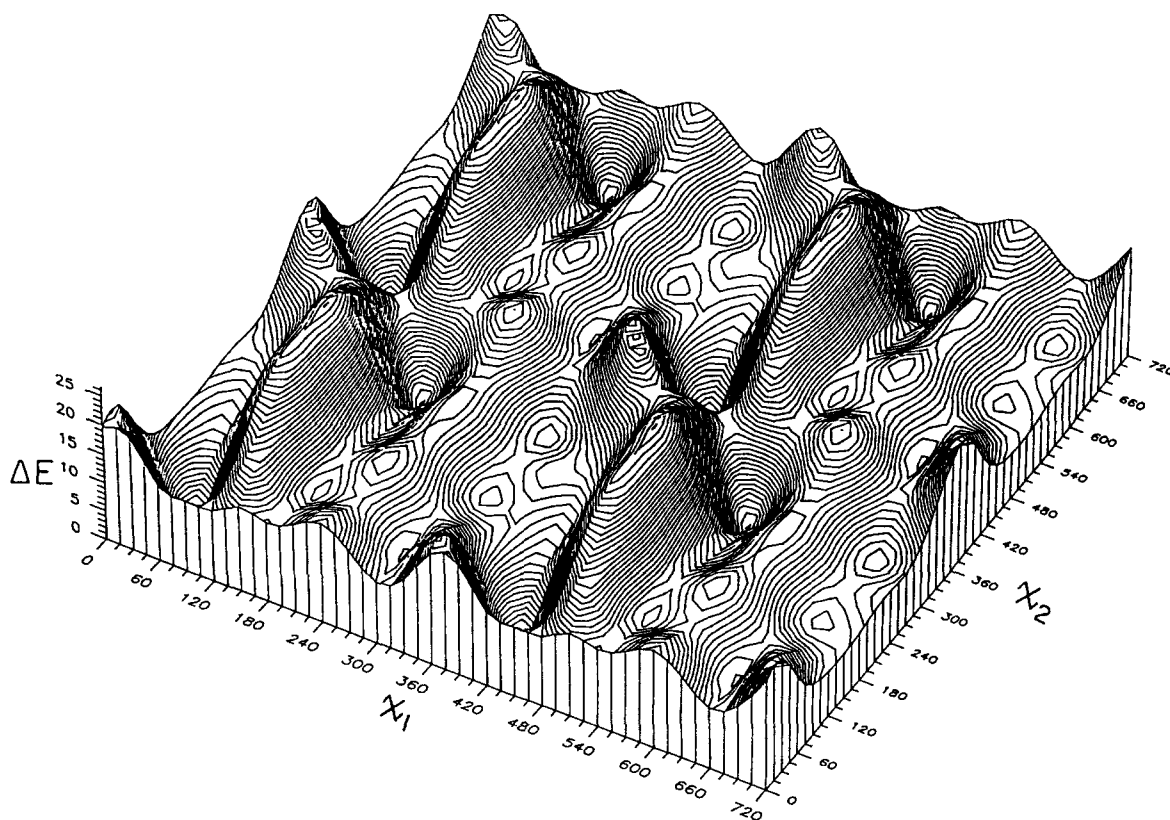


**SCHEME 3.** Backbone ( $\omega$ ,  $\phi$ ,  $\psi$ ,  $\omega'$ ) and side chain ( $\chi_1$ ,  $\chi_2$ ) torsional angles of N-formyl serine amide.

have a maximum of nine different conformations. Three stable orientations can be found about the  $C^\alpha-C^\beta$  ( $\chi_1$  torsional angle), which is to be coupled with an additional three minima arising from the three possible stable orientations about the  $C^\beta-O^\gamma$  ( $\chi_2$  torsional angle) bonds. Using the common *gauche* + ( $g^+$ ), *anti* ( $a$ ), and *gauche* - ( $g^-$ ) description, in an idealized case the locations



**SCHEME 4.** Anticipated minima (A) and maxima (B) of the side chain potential energy surface  $f(\chi_1, \chi_2)$  of serine diamide.



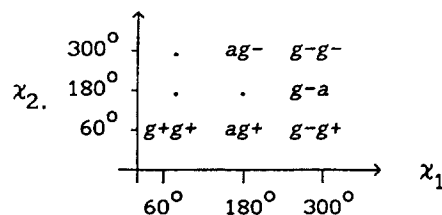
**FIGURE 2.** The rigid sidechain potential energy surface associated with a  $\gamma_L$ -type ( $\phi = -84.5^\circ$ ,  $\psi = +68.7^\circ$ ) backbone orientation ( $f_{\gamma_L}[\chi_1, \chi_2]$ ) of the For-L-Ser-NH<sub>2</sub> molecule, determined by the *ab initio* method (3-21G basis set). In one ( $360^\circ \times 360^\circ$ ) quadrant, 5 maxima, 11 saddle points, and 6 minima were located.

of the nine minima are presumed to be  $g^+$   $g^+$ ;  $[60^\circ, 60^\circ]$ ,  $g^+ a$ ;  $[60^\circ, 180^\circ]$ , ...,  $g^- g^- [-60^\circ, -60^\circ]$ , as shown in Scheme 4A.

On this surface having an idealized topological behavior on the same surface, nine maxima ( $M_A$ ;  $[120^\circ, 0^\circ]$ ,  $2_B$ ;  $[120^\circ, 120^\circ]$ ,  $2_C$ ;  $[120^\circ, 240^\circ]$ , ...,  $M_H$ ;  $[360^\circ, 120^\circ]$ , and  $M_I$ ;  $[360^\circ, 240^\circ]$ ) are also expected (Scheme 4B). Depending on the backbone conformation types,<sup>41,42</sup> due to unfavorable interactions between sidechain/sidechain or sidechain/backbone atoms, the annihilation of some sidechain conformations is expected.

### THE $\gamma_L$ CONFORMATIONS

The sidechain conformation associated with the most stable backbone conformation ( $\gamma_L$ ) was first investigated. For the  $\gamma_L$  backbone conformation, three maxima ( $M_A$ ,  $M_B$ , and  $M_C$ ) are expected to be located around  $\chi_1 = 120^\circ$  on the  $f(\chi_1, \chi_2)$  sur-



**SCHEME 5.** Minima of side chain potential energy surface  $f(\chi_1, \chi_2)$  of  $\gamma_2$  backbone conformation of serine diamide located by *ab initio* computations.

face with an ideal topological character. However, the drastic increase in magnitude of the  $M_B$  maximum located around  $\chi_1 = 120^\circ$ ,  $\chi_2 = 180^\circ$  is the probable source of some topological modifications (Fig. 2.). As a result of a major collapse, the three maxima at  $\chi_1 = 120^\circ$  ( $M_A$ ;  $[120^\circ, 0^\circ]$ ,  $M_B$ ;  $[120^\circ, 120^\circ]$ ,  $M_C$ ;  $[120^\circ, 240^\circ]$ ) became a single maxi-

**TABLE I.** Optimized *ab initio* SCF (3-21G) Geometries of For-L-Ser-NH<sub>2</sub> with  $\gamma_L$ -type Backbone Conformation.

CONV. <sup>a</sup>	$\gamma_L g-g^+$	$\gamma_L g+g^+$	$\gamma_L ag^-$	$\gamma_L ag^+$	$\gamma_L g-g^-$	$\gamma_L g^- a$
$\omega 1$	-178.2	-176.2	-173.8	-177.9	-171.1	-168.8
$\phi 1$	-85.4	-83.6	-83.4	-86.5	-77.4	-77.1
$\psi 1$	+67.4	+71.5	+62.7	+77.8	+63.4	+62.4
$\omega 2$	-179.1	-177.6	-179.9	-176.8	-179.9	-180.0
$\chi 1$	-65.3	+51.9	+179.8	-169.2	-41.2	-44.0
$\chi 2$	+55.2	+69.8	-68.7	+74.8	-75.5	-178.5
O7...HN1	2.09	2.04	2.00	2.16	1.91	1.88
O7...N1	2.93	2.88	2.87	2.96	2.79	2.77
O7.HN1.N1	+140.3	+140.3	+143.4	+135.7	+145.2	+146.1
O3...HO9		2.06	2.03			
O3...O9		2.75	2.72			
O3.HO9.O9		+140.0	126.3			
O9...HN5					2.20	2.22
O9...N5					2.66	2.70
O9.HN5.N5					+106.3	+107.54
MAX. FORCE	<1E-6	<1E-6	<1E-6	<1E-6	<1E-6	<1E-6
E	-486.	-486.	-486.	-486.	-486.	-486.
	.902918	.919567	.911976	.899634	.907605	.907615
$\Delta E$	10.45	0.00	4.76	12.51	7.51	7.50

Torsion angles ( $\omega$ ,  $\phi$ ,  $\psi$ ,  $\chi^1$ , and  $\chi^2$ ) in degrees according to IUPAC-IUB, distances in Angstrom, forces in au, energy (E) in Hartrees, and energy difference ( $\Delta E$ ) in kcal/mol relative to the  $E(\gamma_L[g+g+])[-486.919567$  (Hartree)].

<sup>a</sup> Converged backbone conformation.

mum resulting in the annihilation of three minima ( $g + a$ ,  $g + g -$ , and  $aa$ ) (see Scheme 5).<sup>35</sup>

Due to the modification of the surface topology, minimization started from the ideal ( $g + g -$ ) and ( $g + a$ ) sidechain conformations resulted in the ( $g + g +$ ) conformation, while the ( $aa$ ) initial sidechain geometry converged to the ( $ag -$ ) sidechain conformation (Table I). The observation of five maxima and six minima presume 11 saddle points on the basis of the alternating sum rule.<sup>30,31,43,44</sup> These first-order critical points can be located on the  $f_{\gamma_L}(\chi_1, \chi_2)$  map.

Although only six of the nine minima were found on the  $f_{\gamma_L}(\chi_1, \chi_2)$  surface, all nine ideal sidechain structures

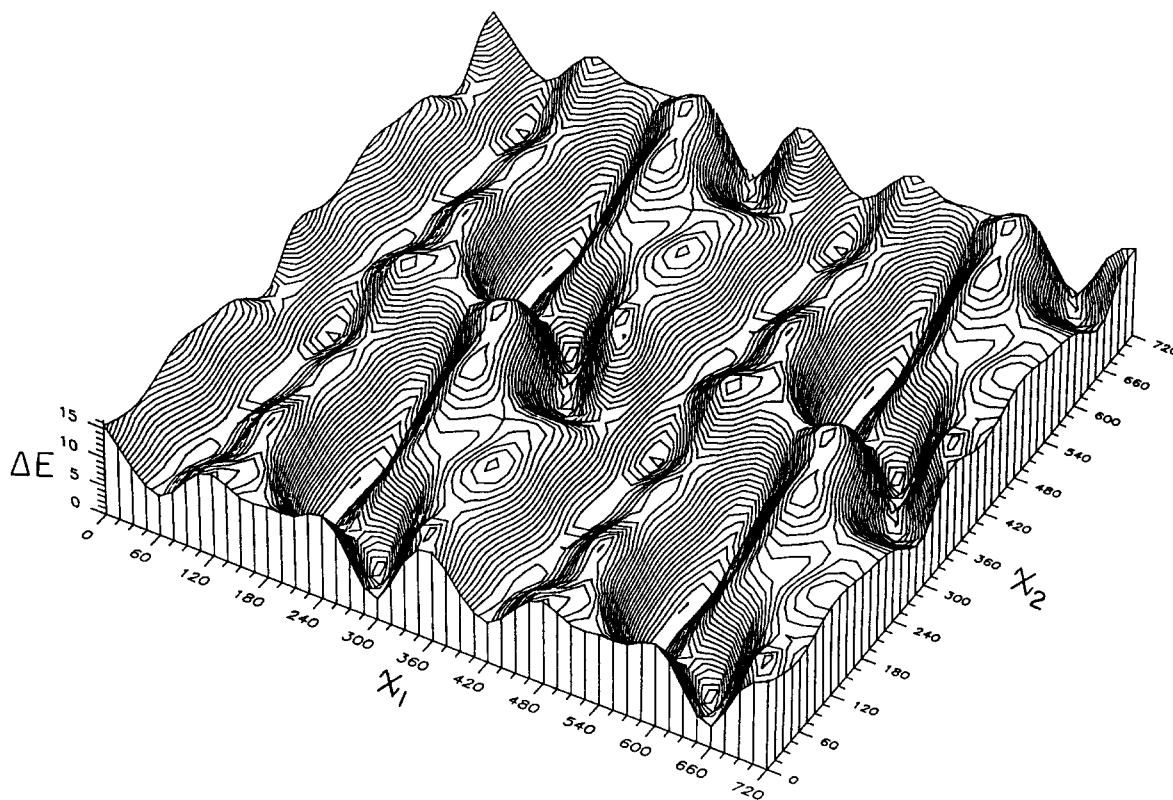
$$([g + g +], [g + a] \cdots [g - g])$$

were submitted for *ab initio* minimization. Selected geometrical properties of the optimized structures are reported in Table I. The  $\gamma_L(g +, g +)$  structure was found to be as the global minimum for the For-L-Ser-NH<sub>2</sub> molecule. Unexpectedly, the

rigid  $f_{\gamma_L}(\chi_1, \chi_2)$  map has a unique feature—namely, further minimization of all the six minima resulted in fully relaxed structures, all of which were located previously on the rigid  $f_{\gamma_L}(\chi_1, \chi_2)$  map. The observation that the  $\gamma_L$ -type backbone minima of the For-L-Ser-NH<sub>2</sub> do no migrate to another backbone type may be related to the strong intramolecular H bond, characteristic of the inverse gamma turn ( $\gamma_L$ ) conformer.<sup>45</sup>

### THE $\beta_L$ CONFORMATIONS

The topological analysis of the  $f_{\beta_L}(\chi_1, \chi_2)$  partially rigid or fixed surface resulted in 6 maxima, 11 saddle points, and 5 minima (Fig. 3), in agreement with the alternating sum rule.<sup>30,31,44,46</sup> The stubby highlands forming a massive mountain block through the right top area (i.e., from the right bottom towards the top left region of the unit surface [ $0^\circ \leq \phi \leq 360^\circ$ ,  $0^\circ \leq \psi \leq 360^\circ$ ]) influence markedly the topology of the surface. The middle section of this mountain block consists of three



**FIGURE 3.** The rigid sidechain potential energy surface associated with a  $\beta_L$ -type backbone ( $\phi = -170.0^\circ$ ,  $\psi = +170.0^\circ$ ) orientation ( $f_{\beta_L}[\chi_1, \chi_2]$ ) of the For-L-Ser-NH<sub>2</sub> molecule was determined by the *ab initio* method (3-21G basis set). In one ( $360^\circ \times 360^\circ$ ) quadrant, 6 maxima, 11 saddle points, and 5 minima were located.

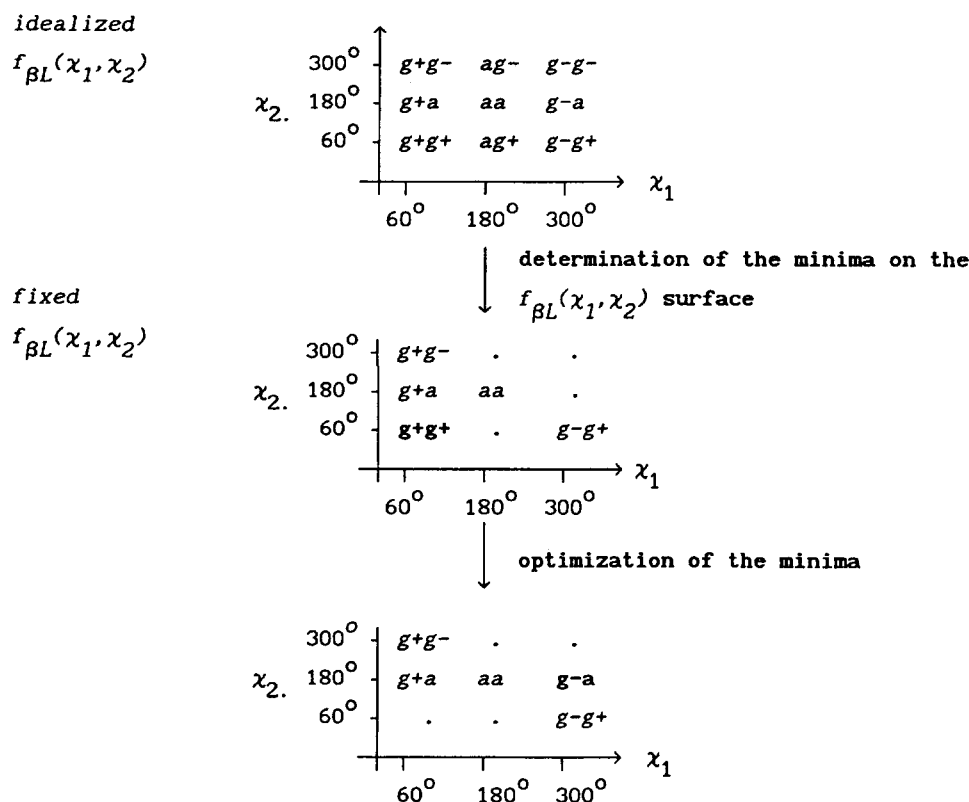
close maxima ( $M_F$ ; [240°, 315°],  $M_E$ ; [250°, 240°], and  $M_H$ ; [330°, 135°]). Consequently, the  $[a, g -]$ ,  $[g -, g -]$ , and  $[a, g +]$  minima were found not to exist. Moreover  $[g +, g +]$  minima located at the left bottom section can be found on the fixed surface; however, on the fully relaxed optimization this minimum also vanished (Scheme 6).

The self-consistent field (SCF) gradient optimization revealed the existence of the  $[g -, a]$  minimum, which was a saddle point on the original surface separating the  $M_E$ ; [250°, 240°] and  $M_H$ ; [330°, 135°] maxima. In fact, this  $[g -, a]$  minimum is associated with a markedly shifted  $\beta_L$  backbone conformation (Table II) as the modified  $\phi$  torsional angle value ( $\phi = -137.3^\circ$ ) is compared to usual  $\beta_L$ -type minima with  $\phi$  value around  $-170^\circ$ , while the value of  $\psi$  was found to be  $+160^\circ$ , only slightly smaller than the normal  $\psi$  value (Table II).

### THE $\delta_L$ CONFORMATIONS

By analyzing the  $f_{\delta L}(\chi_1, \chi_2)$  rigid surface, 5 minima, 10 saddle points, and 5 maxima can be

located (Fig. 4). The high mountain dominating the left side of the surface with a maximum located at  $\chi_1 \approx 150^\circ$ ,  $\chi_2 \approx 210^\circ$  affects the shape of the surface on both sides of the highlands. Four of the six minima— $[g +, g -]$ ,  $[g +, a]$ ,  $[a, a]$ , and  $[a, g +]$ —that are expected to occur in this region of the surface in the case of an ideal topological behavior are in fact eliminated due to the raised mountain ridge. From the optimized *ab initio* grid points associated with a  $\delta_L$ -type backbone conformation, only two (the  $[ag -]$  and the  $[g + g +]$ ) of these six expected sidechain conformations can be identified. One of the two remaining minima, the  $g + g +$  minima, is located in an elongated catchment region (Fig. 4) along the left side of the high maximum. This minimum was assigned as a typical  $[g + g +]$  conformation on the *ab initio* grid surface (Fig. 4), but the total relaxation of the For-L-Ser-NH<sub>2</sub> molecule from the aforementioned initial conformation ( $\delta_L[g + g +]$ ) resulted in a more extended  $\chi_2$  torsional angle value. Consequently, the  $[g + g +] \rightarrow [g + a]$  transition was observed (Scheme 7). On the other hand, the



**SCHEME 6.** Side chain conformational potential energy surface  $f(\chi_1, \chi_2)$  for  $\beta_2$  backbone conformation; *top*: anticipated minima by MDCA; *center*: partially relaxed conformations (with backbone conformation fixed at  $\beta_2$ ); *bottom*: fully relaxed conformations.

**TABLE II.**  
**Optimized *ab initio* SCF (3-21G) Geometries for For-L-Ser-NH<sub>2</sub> with  $\beta_L$ -type Back Bone Conformation.**

CONV. <sup>a</sup>	$\beta_L g-g+$	$\beta_L g+ a$	$\beta_L a a$	$\beta_L g- a$	$\beta_L g+g-$	$\beta_L ag+$
$\omega 1$	176.9	179.7	175.2	-179.5	175.0	
$\phi 1$	-179.0	-170.6	-171.8	-137.3	-166.7	-170.3
$\psi 1$	+172.9	+174.9	-173.4	+160.0	+174.8	-171.5
$\omega 2$	+179.4	+177.7	-179.1	+177.24	+176.6	
$\chi 1$	-89.2	+68.0	-173.0	-64.5	+67.4	-171.8
$\chi 2$	+55.5	-172.9	+155.1	+171.5	-60.5	+90.1
O3...HN5	2.05	2.10	2.06	2.31	2.07	
O3...N5	2.58	2.60	2.58	2.68	2.58	
O3.HN5.N5	+110.7	+108.4	+110.5	+100.1	+109.7	
O7...H09	1.83					
O7...O9	2.71					
O7.H09.O9	+149.0 <sup>o</sup>					
O9...HN1			1.91			
O9...N1			2.69			
O9.HN1.N1			+132.5 <sup>o</sup>			
MAX. FORCE	<1E-6	<1E-6	<1E-6	<1E-6	<1E-6	<1E-6
E	-486.	-486.	-486.	-486.	-486.	-486.
	.902918	.901739	.913531	.895034	.905081	.914309
$\Delta E$	10.45	11.19	3.79	15.39	9.09	3.3

Torsion angles ( $\omega$ ,  $\phi$ ,  $\psi$ ,  $\chi^1$ , and  $\chi^2$ ) in degrees according to IUPAC-IUB, distances in Angstrom, forces in au, energy (E) in Hartrees, and energy difference ( $\Delta E$ ) in kcal/mol relative to the  $E(\gamma_L^2)$  [-486.919567 (Hartree)].

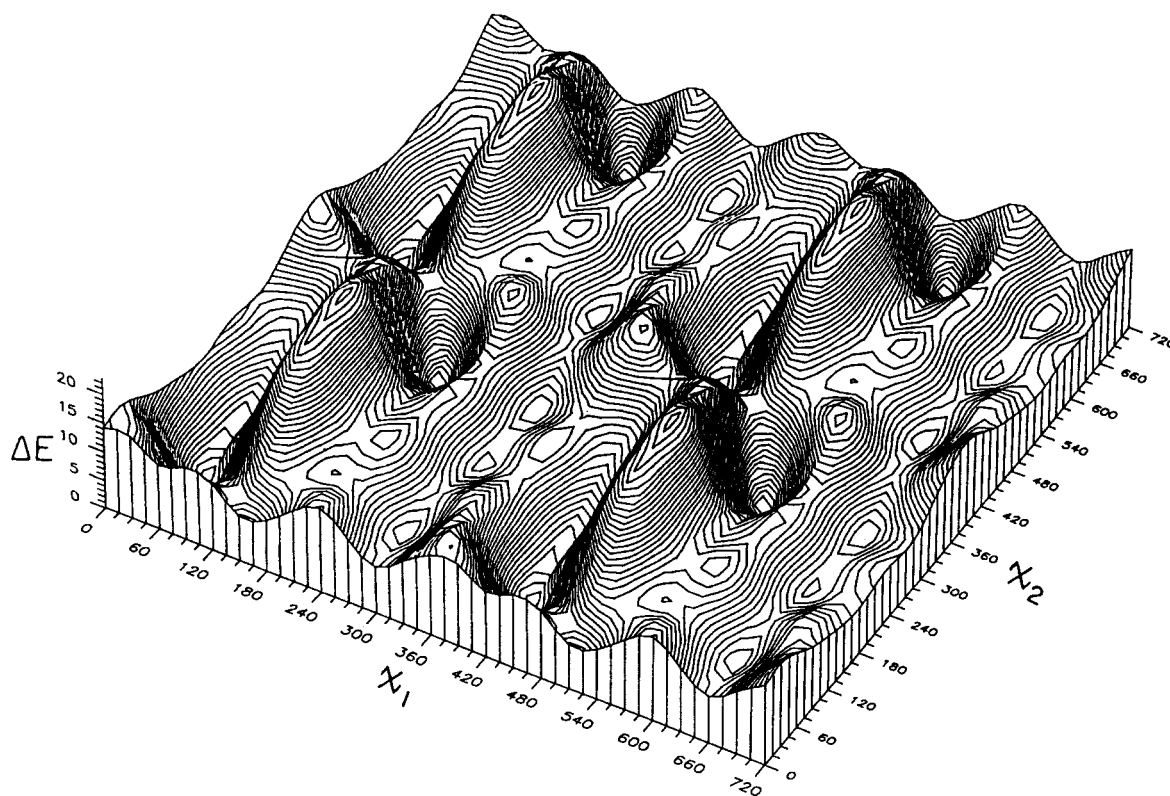
<sup>a</sup> Converged backbone conformation.

$\delta_L$ -type backbone orientation has no major anomaly with the  $[ag-]$ ,  $[g-g-]$ ,  $[g-a]$ , and  $[g-g+]$  sidechain types. Therefore, only minor sidechain shifts from the idealized locations were observed during the computation of the fully relaxed *ab initio* minima (Table III).

Undoubtedly, the type and the relative position of the critical points located on the rigid  $f_{\delta_L}(\chi_1, \chi_2)$  surface (where  $\delta_L[\phi \approx -120^\circ, \psi \approx +30^\circ]$ ) show a great deal of resemblance to the topology of the fixed  $f_{\gamma_L}(\chi_1, \chi_2)$  surface. On both of these two surfaces, the dominating effect of the extremely high mountain incorporating theoretically three maxima ( $M_A$ ;  $[120^\circ, 60^\circ]$ ,  $M_B$ ;  $[120^\circ, 180^\circ]$ , and  $M_C$ ;  $[120^\circ, 300^\circ]$ ) with respect to the idealized surface is the most impressive feature. Furthermore, this

maximum (presumably arising from the collapse of the above three smaller maxima) can be observed on the  $f(\chi_1, \chi_2)$  surfaces associated with the  $\gamma_L$  (ref. 35) as well as with the  $\alpha_D$  (ref. 32) backbone conformations, both of which are neighboring minima of the investigated  $\delta_L$  conformation. Undoubtedly, this similarity is because the  $\delta_L$ ,  $\gamma_L$ , and  $\alpha_D$  backbone conformations have a similar  $\psi$  torsion angle value ( $\psi$  around  $+45 \pm 30^\circ$ ) associated with a different  $\phi$  orientation. In all these cases, due to this observed high maximum, the expected conformations of the For-L-Ser-NH<sub>2</sub> molecule associated with  $\psi \approx +45 \pm 30^\circ$  and  $\chi_1 \approx +120 \pm 45^\circ$  torsional angle combinations are annihilated for all three backbone conformations ( $\delta_L$ ,  $\gamma_L$ , and  $\alpha_D$ ).





**FIGURE 4.** The rigid sidechain potential energy surface associated with a  $\delta_L$ -type backbone ( $\phi = -125.0^\circ$ ,  $\psi = +30.0^\circ$ ) orientation ( $f_{\delta L}[\chi_1, \chi_2]$ ) of the For-L-Ser-NH<sub>2</sub> molecule determined by the *ab initio* method (3-21G basis set). In one ( $360^\circ \times 360^\circ$ ) quadrant, 5 maxima, 10 saddle points, and 5 minima were located.

### THE $\alpha_L$ CONFORMATIONS

Although the nonexistence of the  $\alpha$ -type backbone conformation is reported for other amino acid diamides<sup>13-23</sup> in agreement with MM-type geometries, our preliminary SCF investigation suggested<sup>42</sup> that  $\alpha_L$ -type backbone minima may be present in For-L-Ser-NH<sub>2</sub>. Compared with the appropriate  $\gamma_L$  and/or  $\beta_L$  structures, relatively high conformational energy was predicted for the  $\alpha_L$ -type backbone structures since it fell on the side of a mountain ridge. Computations were performed to explore the stability of the For-L-Ser-NH<sub>2</sub> conformers associated with an  $\alpha_L$ -type backbone orientation in the vicinity of the  $\phi = -60^\circ$ ,  $\psi = -40^\circ$  cross section of the  $\{\phi, \psi, \chi_1, \chi_2\}$  hypercube. As discussed previously,<sup>32</sup> the topological analyses of the  $f_{\alpha L}(\chi_1, \chi_2)$  rigid surface (Fig. 5) revealed 5 maxima, 11 saddle points, and 6 minima. Surprisingly, during the optimization, three of the six minima shown in Scheme 8 migrated to another backbone conformation type upon full geometry optimization. Although on the  $f_{\alpha L}(\chi_1, \chi_2)$

rigid surface the minima and maxima have energy magnitudes within the same relative energy range, the migration of the  $[g+g+]$ ,  $[g+a]$ , and  $[ag--]$  minima were observed.

On the other hand, the remaining three ( $[aa]$ ,  $[g-a]$ , and  $[g-g-]$ )  $\alpha_L$ -type backbone conformations (Table IV) are minima and unique examples that even simple amino acid diamides may adopt the conformational monomer of the right-handed helical secondary structure. The structural features of this outstanding result as well as their minor conformational variations as a function of the applied basis set type are reported in detail elsewhere.<sup>32</sup>

### THE $\epsilon_L$ CONFORMATIONS

The other factor we examined was whether a special serine sidechain conformation could in fact stabilize the second intrinsically unstable conformation,  $\epsilon_L$ . This question is particularly interesting because in the  $\epsilon_L$ -type backbone con-

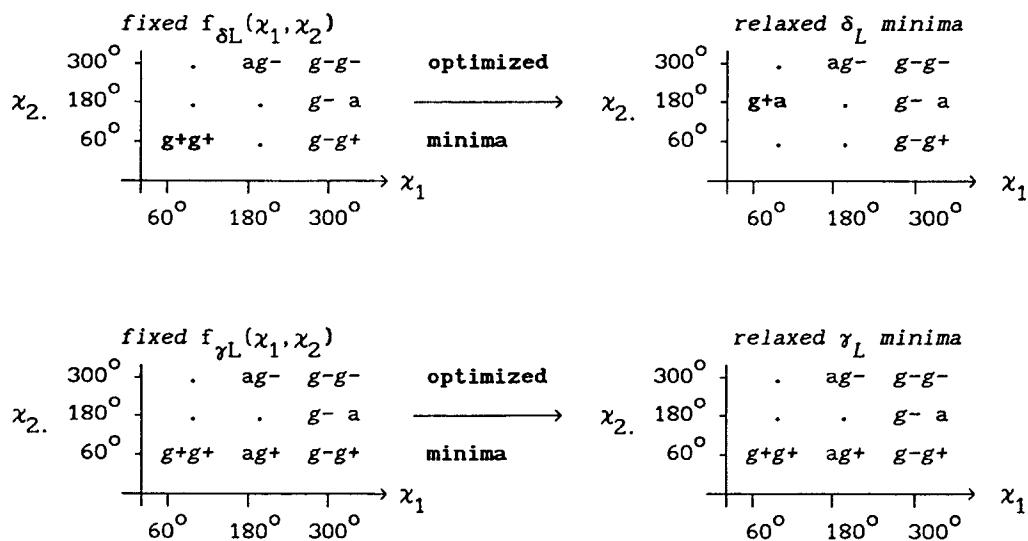
TABLE III.

Optimized *ab initio* SCF (3-21G) Geometries of For-L-Ser-NH<sub>2</sub> with  $\delta_L$ -type Backbone Conformation.

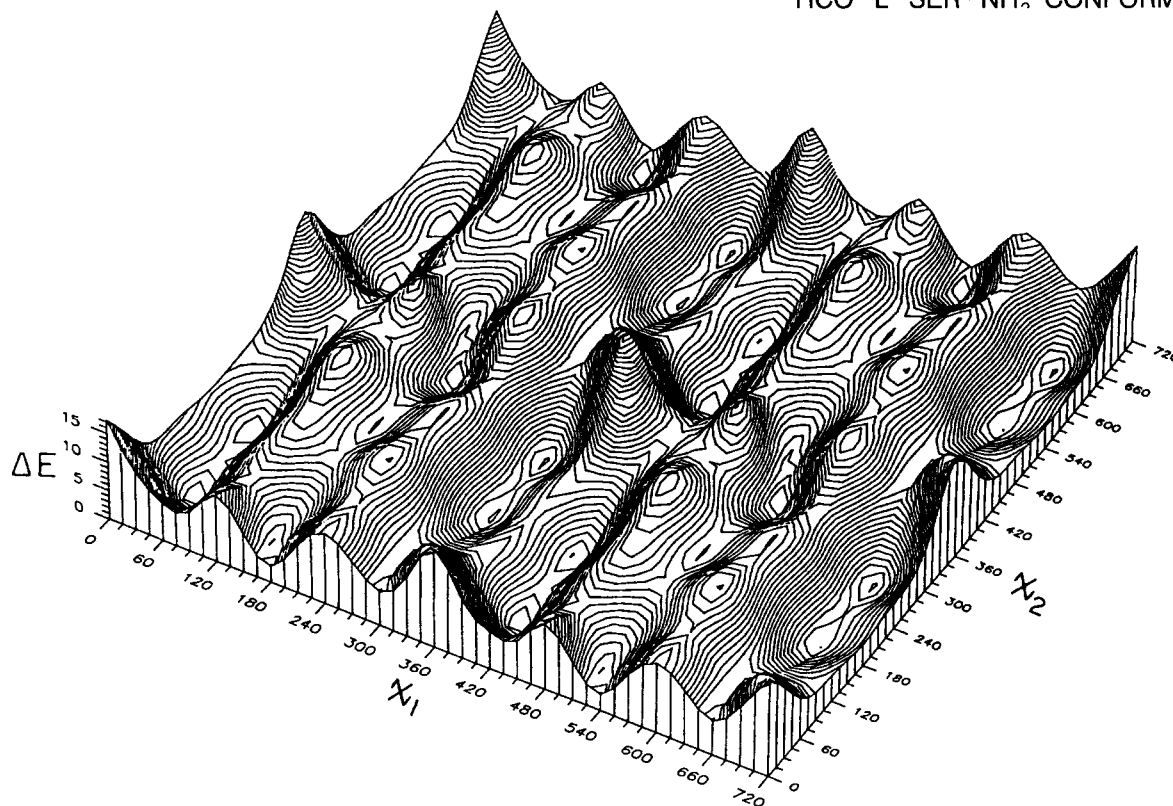
CONV. <sup>a</sup>	$\delta_L g^+ a$	$\delta_L ag^-$	$\delta_L g^- g^-$	$\delta_L g^- a$
$\omega_1$	-172.8	-173.5	-172.0	-166.7
$\phi_1$	-118.1	-128.5	-151.9	-129.9
$\psi_1$	+20.2	+32.9	+35.6	+29.8
$\omega_2$	-177.4	+176.4	+177.0	+176.8
$\chi_1$	+51.4	-172.3	-46.8	-53.1
$\chi_2$	+159.8	-60.5	-54.7	-167.54
O9...HN5	2.19			
O9...N5	2.63			
O9.HN5.N5	+104.7°			
O3...HO9		1.96		
O3...O9		2.69		
O3.HO9.O9		+131.1°		
O7...HO9			2.37	
O7...O9			2.93	
O7.HO9.O9			+116.8°	
MAX. FORCE	<1E-6	<1E-6	<1E-6	<1E-6
E	-486.	-486.	-486.	-486.
	.897209	.906298	.902080	.898408
$\Delta E$	14.03	8.33	10.97	13.28

Torsion angles ( $\omega$ ,  $\phi$ ,  $\psi$ ,  $\chi^1$ , and  $\chi^2$ ) in degrees according to IUPAC-IUB, distances in Angstrom, forces in au, energy (E) in Hartrees, and the energy difference ( $\Delta E$ ) in kcal/mol relative to the  $E(\gamma_L^2)[-486.919567]$  (Hartree).

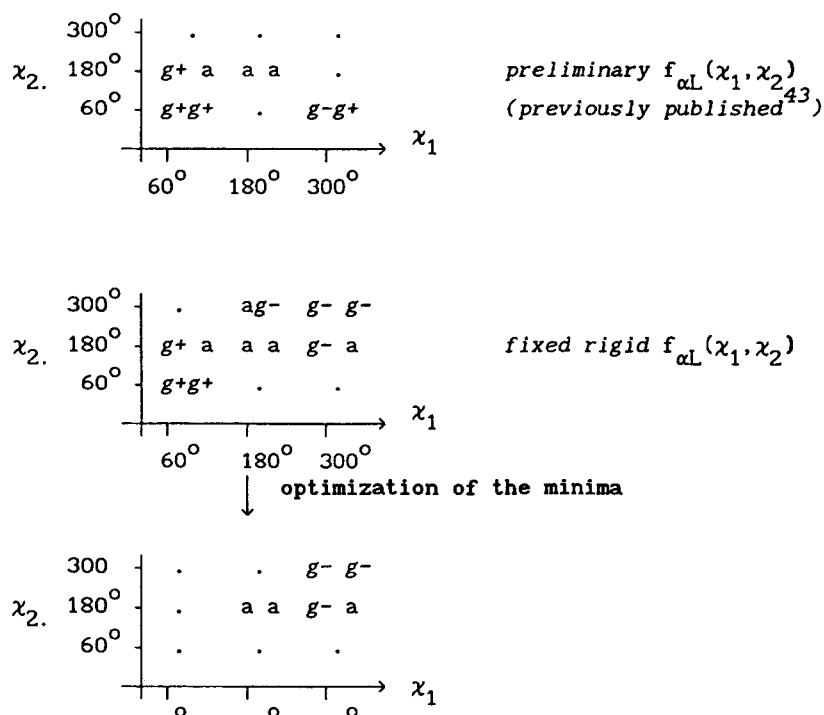
<sup>a</sup> Converged backbone conformation.



**SCHEME 7.** Result of full geometry relaxation on the side chain conformational potential energy surface  $f(\chi_1, \chi_2)$  for the  $\delta_L$  and  $\gamma_L$  backbone conformations.



**FIGURE 5.** The rigid sidechain potential energy surface associated with a  $\alpha_L$ -type backbone ( $\phi = -60.0^\circ$ ,  $\psi = -40.0^\circ$ ) orientation ( $f_{\alpha_L}[\chi_1, \chi_2]$ ) of the For-L-Ser-NH<sub>2</sub> molecule determined by the *ab initio* method (3-21G basis set). In one ( $360^\circ \times 360^\circ$ ) quadrant, 5 maxima, 11 saddle points, and 6 minima were located.



**SCHEME 8.** Side chain conformational potential energy surface  $f(\chi_1, \chi_2)$  for  $\alpha_2$  backbone conformation; *top*: results of a preliminary optimization;<sup>42</sup> *center*: partially relaxed conformations (with backbone conformation fixed at  $\alpha_L$ ); *bottom*: fully relaxed conformations.

**TABLE IV.**  
Optimized *ab initio* Structures of For-L-Ser-NH<sub>2</sub>  
at  $\alpha_L$ -type Backbone Conformation, Using 3-21G  
Basis Set.

CONV. <sup>a</sup>	$\alpha_L g^- g^-$	$\alpha_L g^- a$	$\alpha_L a a$
$\omega 1$	-171.1	-169.4	-174.6
$\phi 1$	-72.0	-70.5	-62.4
$\psi 1$	-23.7	-24.9	-42.8
$\omega 2$	157.0	-179.8	-179.7
$\chi 1$	-45.6	-47.5	+179.9
$\chi 2$	-77.5	-174.5	-168.7
O9...HN5	2.21	2.14	
O9...N5	2.70	2.63	
O9.HN5.N5	108.4	108.3 <sup>o</sup>	
N1 out.	-22.0	-25.5	-12.6
N2 out.	6.8	7.9	11.1
MAX. FORCE	<1E-6	<1E-6	<1E-6
E	-486.	-486.	-486.
	.8996122	.892646	.8867149
$\Delta E$	12.5	16.9	20.6

Torsion angles ( $\omega$ ,  $\phi$ ,  $\psi$ ,  $\chi^1$ , and  $\chi^2$ ) in degrees according to IUPAC-IUB, distances in Angstrom, forces in au, energy (E) in Hartrees, and energy difference ( $\Delta E$  absolute) in kcal/mol relative to the  $E(\gamma_L)$  [-486.919567 (Hartree)].

<sup>a</sup> Converged backbone conformation.

formations, there is no structural possibility for intramolecular Bronsted complex formation between sidechain and backbone atoms. The rigid sidechain potential energy surface associated with an  $\varepsilon_L$ -type backbone ( $\phi = -95.0^\circ$ ,  $\psi = +150.0^\circ$ ) orientation ( $f_{\varepsilon_L}[\chi_1, \chi_2]$ ) of the For-L-Ser-NH<sub>2</sub> molecule determined by the *ab initio* method (3-21G basis set) is shown in Figure 6. In one quadrant, 8 maxima, 15 saddle points, and 7 minima were located. *Ab initio* calculations were performed on all possible sidechain conformations

$$[g^+g^+], [g^+a], [aa], \dots, [g^-g^-].$$

Typically, the sidechain conformation remained unchanged, but the backbone conformation migrated to one of the nearest neighbors of the  $\varepsilon_L$

backbone conformer—namely, to  $\beta_L$  or  $\gamma_L$  structures, as shown in Scheme 9.

## Conclusion

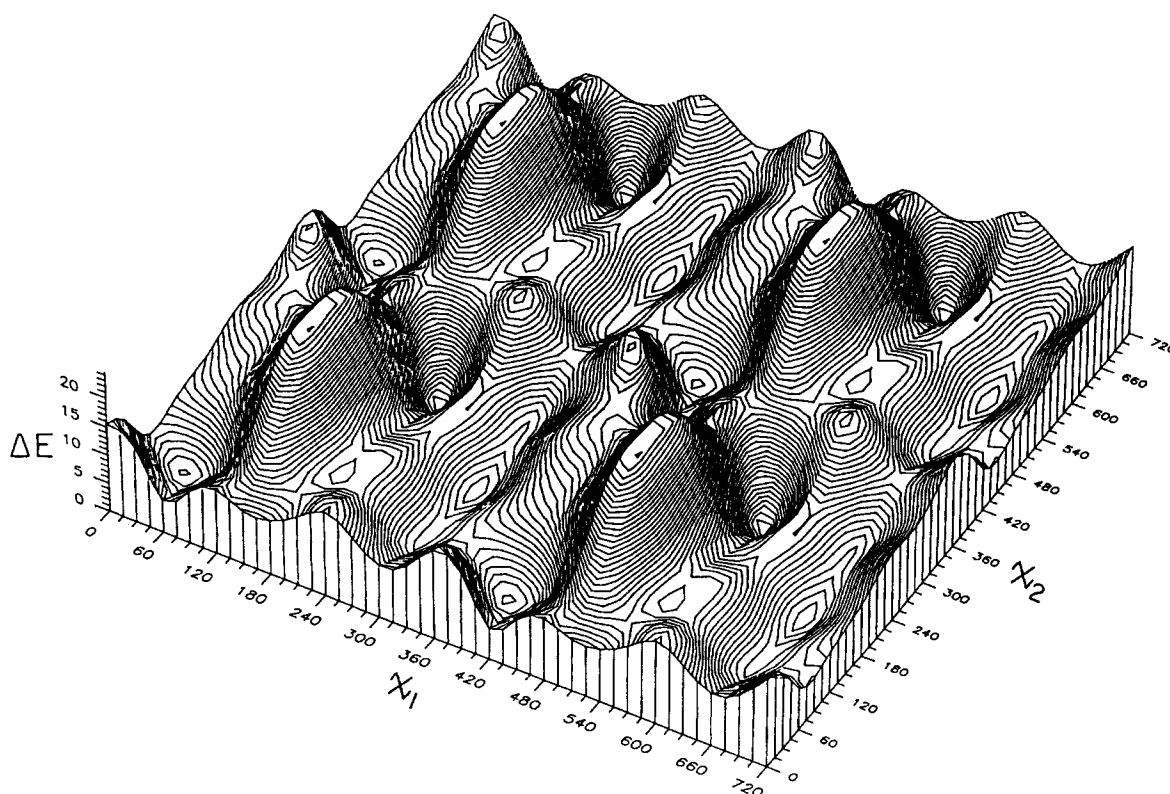
Using the multidimensional conformation analysis<sup>30,31</sup> of Ramachandran-type potential energy surfaces, it was previously demonstrated<sup>21,22</sup> that for simple amino acid diamides a maximum of nine legitimate backbone minima ( $\alpha_L$ ,  $\alpha_D$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\gamma_D$ ,  $\varepsilon_L$ ,  $\varepsilon_D$ ,  $\delta_L$ ,  $\delta_D$ ) may occur. In the case of the amino acid diamide, such as the serine diamide with two sidechain torsional angles ( $\chi_1$ ,  $\chi_2$ ), not less than  $3 \times 3 = 9$  minima are theoretically presumed. These nine sidechain conformers may be associated with each of the previously mentioned typical backbone conformations. Since for diamides with an L (C<sup>α</sup>) configuration the stability of the  $\alpha_L$ -type backbone conformers is expected to be lower in energy than the  $\alpha_D$ -type structures, the systematic sidechain conformational mapping with  $\alpha_L$ ,  $\beta_L$ ,  $\gamma_L$ ,  $\delta_L$ , and  $\varepsilon_L$  backbone conformations was expected to reveal the most important structures of

For-L-Ser-NH<sub>2</sub>.

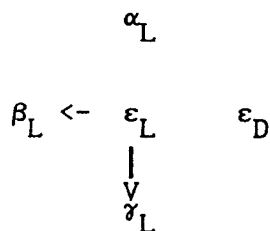
To reduce the possibility of skipping over minima with high energy or to miss structures located in hidden conformation valleys, five complete grids (a total of  $5 \times 169 = 845$  *ab initio* grid points) were determined. As a result, a total of  $6\gamma_L$ ,  $6\beta_L$ ,  $4\delta_L$ ,  $3\alpha_L$ , and no  $\varepsilon_L$  fully relaxed structures (a total of 19 conformers) were located for the For-L-Ser-NH<sub>2</sub> molecule, using *ab initio* type (3-21G) calculations. However, no  $\varepsilon_L$  conformation has been found.

## Acknowledgments

This research was supported by grants from the Hungarian Scientific Research Foundation (OTKA F013799 and T017192). Thanks are also due to the Higher Education Development Fund (FEFA) for providing computational facilities (CFA/III-254). The continued financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada is gratefully acknowledged. The helpful technical assistance of R. Pálvölgyi is appreciated.



**FIGURE 6.** The rigid side chain potential energy surface associated with a  $\epsilon_L$ -type backbone ( $\phi = -95.0^\circ$ ,  $\psi = +150.0^\circ$ ) orientation ( $f_{\epsilon_L}[\chi_1, \chi_2]$ ) of the For-L-Ser-NH<sub>2</sub> molecule determined by the *ab initio* method (3-21G basis set). In one ( $360^\circ \times 360^\circ$ ) quadrant, 8 maxima, 15 saddle points, and 7 minima were located.



**SCHEME 9.** Backbone conformational shift to one of the nearest neighbors from the initial  $\epsilon_L$  backbone structure coupled with any possible side chain conformation during geometry optimization.

## References

1. K. Wüthrich, *NMR of Protein and Nucleic Acids*, John Wiley & Sons, New York, 1986.
2. H. J. Dyson, M. Rance, R. A. Houghten, R. A. Lerner, and P. E. Wright, *J. Mol. Biol.*, **201**, 161 (1988).
3. H. J. Dayson and P. E. Wright, *Ann. Rev. Biophys. Chem.*, **20**, 519 (1991).
4. S. J. Stradley, J. Rizo, M. D. Bruch, A. N. Stroup, and L. M. Gierasch, *Biopolymers*, **29**, 263 (1990).
5. A. Perzcel, M. Hollósi, B. M. Foxman, and G. D. Fasman, *J. Am. Chem. Soc.*, **113**, 9772 (1991).
6. D. Neuhaus and M. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH Publications, Verlagsgesellschaft, 1989.
7. M. Vásquez and H. A. Scheraga, *J. Biomol. Struct. Dynamics*, **5**, 705 (1988).
8. F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
9. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **107**, 3902 (1985).
10. S. J. Weiner, U. C. Singh, T. J. O'Donnell, and P. A. Kollman, *J. Am. Chem. Soc.*, **106**, 6243 (1984).
11. B. B. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, **4**, 187 (1983).
12. F. J. Seiler, Res. Lab of the U.S. Air Force Academy, Mopac 5.00, QCPE No. 455, Colorado Springs, CO, 1988.
13. H. L. Sellers and L. Schäfer, *J. Am. Chem. Soc.*, **100**, 7728 (1978).
14. L. Schäfer, H. L. Sellers, F. J. Lovas, and R. D. Suenram, *J. Am. Chem. Soc.*, **102**, 6566 (1980).
15. L. Schäfer, C. Van Alsenoy, and J. N. Scarsdale, *J. Chem. Phys.*, **76**, 1439 (1982).
16. V. J. Klimowski, L. Schäfer, F. A. Momany, and C. Van Alsenoy, *J. Mol. Struct. (THEOCHEM)*, **124**, 143 (1985).

17. J. N. Scarsdale, C. Van Alsenoy, V. J. Klimkowski, L. Schäfer, and F. A. Momany, *J. Am. Chem. Soc.*, **105**, 3438 (1983).
18. L. Schäfer, V. J. Klimkowski, F. A. Momany, H. Chuman, and C. Van Alsenoy, *Biopolymers*, **23**, 2335 (1984).
19. T. Head Gordon, M. Head Gordon, M. J. Frish, C. Brooks II, and J. A. Pople, *Int. J. Quantum Chem. Quantum Biology Symposium*, **16**, 311 (1989).
20. T. Head Gordon, M. Head Gordon, M. J. Frish, C. Brooks II, and J. A. Pople, *J. Am. Chem. Soc.*, **113**, 5989 (1991).
21. A. Perczel, J. G. Ángyán, M. Kajtár, W. Viviani, J.-L. Rivail, J.-F. Marcoccia, and I. G. Csizmadia, *J. Am. Chem. Soc.*, **113**, 6256 (1991).
22. M. A. McAllister, A. Perczel, P. Császár, W. Viviani, J.-L. Rivail, and I. G. Csizmadia, *J. Mol. Struct. (THEOCHEM)*, **288**, 161 (1993).
23. W. Vladia, J.-L. Rivail, A. Perczel, and I. G. Csizmadia, *J. Am. Chem. Soc.*, **115**, 8321 (1993).
24. M. A. McAllister, A. Perczel, P. Császár, and I. G. Csizmadia, *J. Mol. Struct. (THEOCHEM)*, **288**, 181 (1993).
25. A. Perczel, M. A. McAllister, P. Császár, and I. G. Csizmadia, *J. Am. Chem. Soc.*, **115**, 4849 (1993).
26. C.-M. Liegener, G. Endredi, M. A. McAllister, A. Perczel, J. Ladik, and I. G. Csizmadia, *J. Am. Chem. Soc.*, **115**, 8275 (1993).
27. A. Perczel and I. G. Csizmadia, *J. Mol. Struct. (THEOCHEM)*, **286**, 75 (1993).
28. A. Perczel, M. A. McAllister, P. Császár, and I. G. Csizmadia, *Can. J. Chem.*, **72**, 2050 (1994).
29. M. Cheung, M. E. McGovern, T. J. Danchman, T. Jin, D. C. Zhao, M. A. McAllister, Ö. Farkas, A. Perczel, P. Császár, and I. G. Csizmadia, *J. Mol. Struct.*, **309**, 151 (1994).
30. I. G. Csizmadia, In *The Chemistry of the Thiol Group*, a member of the series *The Chemistry of Functional Groups*, Saul Patai, Ed., John Wiley & Sons, 1974, p. 1. (See particularly pages 36–41, including Figures 23 and 23 as well as Table 20.)
31. I. G. Csizmadia, In *New Theoretical Concept for Understanding Organic Reactions*, J. Bertrán, Ed., D. Reidel Publishing Co., Dordrecht, 1989, p. 1.
32. A. Perczel, Ö. Farkas, J. F. Marcoccia, and I. G. Csizmadia, *Int. J. Quant. Chem.*, submitted.
33. G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan, *J. Mol. Biol.*, **7**, 95 (1963).
34. (a) IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970); (b) F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, Jr., J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, *J. Mol. Biol.*, **112**, 535 (1977); (c), In *Data Commission of the International Union of Crystallography*, E. E. Abola, F. G. Bernstein, S. H. Bryant, I. F. Koetzle, and J. Weng, Eds., Bonn/Cambridge/Chester, 1987, p. 107.
35. Ö. Farkas, A. Perczel, J. F. Marcoccia, M. Hollósi, and I. G. Csizmadia, *J. Mol. Struct. (THEOCHEM)*, **331**, 27 (1995).
36. P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, *J. Am. Chem. Soc.*, **101**, 2550 (1979).
37. G. Fogarasi, X. Zhou, P. W. Taylor, and P. Pulay, *J. Am. Chem. Soc.*, **114**, 8191 (1992).
38. P. Császár and P. Pulay, *J. Mol. Struct. (THEOCHEM)*, **114**, 31 (1984).
39. J. Baker, *J. Comput. Chem.*, **7**, 385 (1986).
40. (a) P. Pulay and Associates, TX90, Fayetteville, AR, 1990; (b) P. Pulay, *Theor. Chim. Acta*, **50**, 229 (1979); (c) J. Frank, Selier Res Lab., U.S. Air Force Academy, Mopac 5.00 QCPE No. 455, Colorado Springs, CO, 1988; (d) M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **107**, 3902 (1985).
41. S. S. Zimmerman, M. S. Pottle, G. Nemethy, and H. A. Scheraga, *Macromolecules*, **10**, 1 (1977).
42. A. Perczel, R. Daudel, J. G. Ángyán, and I. G. Csizmadia, *Can. J. Chem.*, **68**, 1182 (1990).
43. In *Progress of Theoretical Organic Chemistry*, M. R. Peterson and I. G. Csizmadia, I. G. Csizmadia, Eds., Elsevier, Amsterdam, Vol. 3, 1990, p. 190.
44. P. G. Mezey, *Potential Energy Hypersurfaces*, Elsevier Science Publishers, Amsterdam, 1987.
45. J. A. Smith and L. G. Pease, In *CRC (Chemical Rubber Company) Critical Reviews in Biochemistry*, Vol. 8, 1980, p. 315, Gerald D. Fasman Ed.
46. J. G. Ángyán, R. Daudel, Á. Kucsman, and I. G. Csizmadia, *Chem. Phys. Lett.*, **136**, 1 (1987).